COPYRIGHT NOTICE

Title: Level of agreement between field-based data collectors in a large scale injury prevention randomised controlled trial

Running Head: Field-based data collection methods

Dara M. Twomey[1], Caroline F. Finch[1], Tim L.A. Doyle[2, 3], Bruce C. Elliott[2], David G. Lloyd[2]

[1]School of Human Movement and Sport Sciences, University of Ballarat, Victoria, Australia.

[2]School of Sport Science, Exercise and Health, University of Western Australia, Crawley, Western Australia, Australia.

[3]School of Medical Sciences, RMIT University, Melbourne, Victoria, Australia.

Corresponding Author:

Dara Twomey

School of Human Movement & Sport Sciences

University of Ballarat,

PO Box 663

Ballarat

VIC 3353

Australia

d.twomey@ballarat.edu.au

Acknowledgements

The PAFIX study was funded by a nationally-competitive research grant from the (Australian) National Health and Medical Research Council (NH&MRC) – Project ID 400937. DT and TD were funded by research fellowships through this grant and CFF was supported by an NHMRC Principal Research Fellowship. DL also received funding from the Western Australian Medical and Health Research Infrastructure Fund.

Abstract

In sports injury prevention field trials, data collectors are often club volunteers with considerable knowledge of the game but with limited detailed medical backgrounds or knowledge of formal scientific processes. The aim of this paper is to determine the agreement among trained primary data collectors (PDCs) with a sport science background and no prior involvement in data collection in a large randomised controlled trial. During the 'Preventing Australian Football Injury through eXercise' (PAFIX) project, player participation and injury data were collected by trained PDCs at training and games over the 2007 and 2008 playing seasons in 40 community level Australian football teams. PDC-collected data relating to player exposure and whether or not a player sustained an injury and subsequently left the field of play was compared to the same information from independent observers (IOs) who attended one randomly selected game for each of the 40 teams. There was 98% agreement between the PDC and the IO on game details (i. e., date, time, grade and score), 79% (ICC 0.9, 95%CI 0.85–0.95) agreement on the number of players per game and 76% (ICC 0.8, 95%CI 0.69–0.91) agreement on the number of injuries sustained in the games. There was 100% agreement on whether the player left the field for all injuries. This study found that exercise and sport science students are reliable data collectors in sports injury fieldwork studies.

Keywords= Data collection tools, injury prevention, agreement, randomised controlled trial

Introduction

Accurate and reliable data are critical to the success and validity of any research project. In sports injury prevention field trials, important data pertaining to the number and types of injuries sustained and exposure to injury risk are collected by data collectors, who are often volunteers or club/team members with considerable knowledge of the game and players but with limited medical backgrounds or knowledge of formal scientific processes. Therefore, validation of data collection methods and injury incidence statistics is critical for ensuring that data used for the calculation of injury rates are accurate. As injury rates are broadly defined as the number of injuries divided by the amount of exposure to risk, it is highly important that both the numerator and denominator are accurate.

Although many different methods of data collection have been used in sports injury studies, few have examined, reported or acknowledged the data quality from those approaches.[1-4] A major characteristic of club/team based injury surveillance studies is that a designated data collector records details of injuries that occur during a game and/or training session on-site, or within a very short timeframe after their occurrence. Accordingly, some studies have used sports trainers/first aiders or other club volunteers whilst others have used medical support staff to collect injury and/or exposure data. For example, in the context of non-elite Australian football injury surveillance studies, data collectors have ranged from physiotherapists [5-7], sports trainers/first aiders [8] to formally trained primary data collectors (PDCs).[9-12] The extent to which injuries can be adequately described and correctly assigned with reference to nature, body region and causation depends upon the training and experience of the data collector.[2, 4]

Data collection procedures relying on medical or health care professionals can be costly and whilst they often exist at elite level, this is not usually the case for community-based sport.[3] Therefore, due to logistical and financial constraints, large scale community-based cohort

studies frequently use PDCs to collect injury and exposure data. However, irrespective of the background of the PDC, inaccurate and unreliable data will have serious implications in the calculation of accurate injury rates and may lead to inappropriate preventive strategies.

Although most injury prevention studies describe their injury data collection methods, rarely has the accuracy of these processes been reported in terms of case ascertainment.[13-15] Their level of accuracy can be determined by comparison: against an external independent source of high quality injury data or by comparison with independent observation of the injury events. In some injury surveillance studies, it has been possible to validate the injury description and severity against the medical reports from the hospital admission data or health professional records. For example, a junior community Australian football injury study, found only minor discrepancies existed upon verifying data collected by club based data collectors and formal medical records.[16] A study of protective equipment use in community level Australian football, verified the PDCs' data by having an independent observer attend the same games to record the same information.[11] Although slight differences existed between the PDC recorded exposure data and personal protective equipment use and the audits completed by research staff, these differences were non-significant. This verification approach appears to be appropriate for assessing the agreement between formal exposure data and data that are identifiable from the sideline of the playing area in team sports. In that study [11], and most previous studies using PDCs [9-12], data were collected by volunteers from the club with an extensive knowledge of the players and game.

To date, no study has assessed the use of trained PDCs with a sports science background and an understanding of research processes in the collection of injury and related exposure data. The aim of this paper is to determine the reliability of trained and remunerated PDCs with a sport science background and no prior involvement with the relevant sports teams, for injury and participation data collection in a large randomised controlled trial (RCT) conducted in community level Australian football.

Methods

During the 'Preventing Australian Football Injury through eXercise' (PAFIX)[17] RCT, player participation, injury and ground condition data were collected by trained PDCs at all training sessions and games over the 2007 and 2008 playing seasons in 18 community level Australian football clubs. This included a total of 40 teams ranging from senior, reserve and colt (under 19) grades in 10 clubs in Victoria and 8 clubs in Western Australia over the two years. Institutional ethics approval was obtained prior to the commencement of the study and full details of the RCT protocol are published elsewhere.[17]

The PDCs were selected from undergraduate and graduate exercise and sports science students at the universities managing the PAFIX project. Selection was based on: a) experience or keen interest in injury prevention; b) a commitment to attend all training sessions and games each week; and c) good communication skills. A total of 20 PDCs were employed each year and one was assigned to each of the participating PAFIX teams. All PDCs attended a two day training session, which included familiarisation with the PAFIX training programs and data collection procedures. A detailed procedure manual was also provided to each PDC as a reference guide. To ensure consistency in data collection procedures, this training was delivered by the same researcher across the two states. All completed exposure, injury and ground condition forms were collected weekly. Scrutiny of forms ensured a follow-up occurred by the project team when data were incomplete or required clarification.

As a data collection audit process, independent observers (IOs) also attended one randomly selected game for each of the 40 PAFIX teams over each 18-week playing season. On two occasions, project teams were playing against each other and data for both teams were collected, giving a total of 42 agreement assessments. The four IOs, two each season from each state, were selected from the same group as the PDCs and the same IOs were used to

record data for all the games during the season of their recruitment. The structure of the data collection and importance of accuracy was stressed by the same researcher who delivered the PDC training. They were also provided with written instructions on completing the standardised data collection form (Supplementary A). They were explicitly requested to remain inconspicuous to the regular PDC at the game. For the validation exercise, data were collected on game details, such as date, time, grade and score, player participation and whether or not a player sustained an injury and subsequently left or remained on the field of play. Because the IOs were unknown to the team, they collected player data based solely on player jumper number. In contrast, the PDCs completed both player name (as required for follow-up during the RCT) and player jumper number. The definition of injury used (i.e. something that caused the player to seek medical attention or to leave the field of play), was reiterated on the IO instruction sheet. Validation of specific injury details was not undertaken in this study as the IOs did not have access to the sports trainers, medical support staff or players to obtain this information.

Data were double entered in Excel and converted to SPSS^{TM} (version 17.0) for analyses. Tabular frequencies and percentage agreements were calculated for all participation and injury data. Intra class correlations (ICC) with 95% confidence intervals (CI) were calculated for all discrete numerical data. The ICCs were rated against the following scale of acceptance: almost perfect agreement 0.81-0.99, substantial agreement 0.61–0.80, moderate agreement 0.41–0.60, fair agreement 0.21–0.40, and slight agreement 0.01–0.20.[18]

Results

There was effectively total agreement (98%) between the PDC and the IO on all the game details including date, time, grade and score. On one occasion a PDC recorded a win and

the IO recorded a loss, however, the PDC's score was determined to be correct according to competition coordinators.

When the participation exposure data were examined by the number of players per game, there were 33/42 (79%) games that had 100% agreement (Table 1). Variations in the other nine games were almost equally divided between the PDCs and IOs (IOs listed one additional player on five occasions and the PDCs listed one additional player on three occasions and two additional players on one occasion).

<Insert Table 1 here>

On initial analysis, the level of agreement for the number of game injuries was relatively poor at 52%. On 21 occasions the same number of injuries was recorded; on seven occasions the PDC noted more and on 14 occasions the IO recorded more (Table 2). The difference in the number of injuries between the PDC and IO was predominantly one injury, with a difference of two recorded on three occasions and a difference of four injuries once. On closer examination of the data, it was evident that on five out of the seven times where the PDC recorded more injuries, the injury occurred towards the end of a quarter and the injury was treated during the break. When adjusted for these injuries, as it was not possible for the IO to collect this information, the level of agreement increased to 76%. Similarly, the level of agreement regarding the number of injured players who left the field increased when adjusted for the games where both the PDC and IO recorded an injury. This adjustment resulted in total agreement between the PDC and IO for the players who left the field.

<Insert Table 2 here>

Discussion

The results of this study highlight the importance of reporting the reliability of data collection methods to validate the integrity of research data to facilitate correct interpretation of estimated injury rates. In any situation where match and training data underpin the development or evaluation of injury prevention strategies, the accuracy of the results is critical.[2, 19, 20]

While there was substantial agreement (79%) on the number of players who participated in each game between the PDCs and IOs, the absolute value is lower than previously reported in a similar game context.[19] Although perfect agreement on this variable might be expected, players switching jumpers was reported by a PDC on one occasion and if this occurred unnoticed in other matches, then substantial agreement may be acceptable.

As exposure estimates are a key element in the calculation of injury rates, a high level of agreement is essential for accurate injury rates. Ideally, the PDC data would be correct but as only two sources of information were compared, no firm conclusions can be drawn. If, in fact, it was the IOs who were fully correct in their assessments, then the PDCs had reporting agreements that were 76% and 79% of those of the IO on number of injuries and player participation respectively. The true situation is likely to lie in between, with the PDCs correctly reporting within the range of 76-100% for injury numbers and 79-100% agreement for the participation data. It is very likely that many other studies also suffer from underreporting of injuries and player exposures. However, unless they report the accuracy of their data, it is difficult to establish whether they have substantially under or overestimated their injury rates and hence whether the rates are truly comparable with other studies.

Player participation was recorded in terms of numbers of players who participated in the game and the high level of agreement indicates that this data provide acceptable figures for the calculation of general exposure estimates. However, if more specific exposure estimates,

such as per game quarter, were required for any study outcomes then, the agreement level of the more specific time intervals would need to be included in the agreement assessment.

The initial level of agreement on the number of injuries sustained was unexpectedly low considering the clarity of the injury definition and level of training provided to both the PDCs and IOs. The substantial increase in the level of agreement following fine scrutiny of the data, where variations were evident, highlights the need for this process. Almost all injuries that were reported by the PDC, but not the IO, were identified and treated outside actual game play. The nature of the agreement assessments did not facilitate the IOs to identify and report these injuries, as they were requested to remain inconspicuous and not engage with the PDC, team medical staff or the players at any time. This is an important consideration in reporting accurate data in similar future studies. The high level of agreement on the severity indicator included in this study (i.e., whether or not the injured player left the playing field) was classified as 'perfect' and demonstrated that these data were reported accurately.

Most PDCs used in previous Australian football injury studies [12, 19] have been volunteers from within the clubs with an extensive knowledge of the game and the players and a very strong commitment to the club environment. In this study, the PDCs were selected from exercise and sports science courses and therefore had a basic knowledge of the musculoskeletal system but no prior engagement with the club. The PDCs' fundamental knowledge of the human body and the appreciation for scientific rigour was a valuable asset in identifying and reporting the region and nature of all injuries. The PDCs were also remunerated for the data collection which is not always possible within budgetary constraints and limited funding. However, most university courses in exercise and sport science have practical and professional placement requirements for students and could be considered as a viable source for PDCs in future injury prevention studies.

A limitation of this study is that there was no gold standard to compare either data set. An attempt was made to video the games where the validation exercise was being undertaken, however, the size of the Australian football field made it too difficult to distinguish the player jumper numbers of the injured players. With constant developments in technology, this may be possible in future studies.

Conclusion

In conclusion, the results highlight the importance of undertaking data collection audits as part of the research design. Moreover, it would seem that exercise and sport science students are reliable data collectors in sports injury projects in relation to both player game exposure and injury numbers. Reporting the accuracy of results (both in terms of reliability and validity) should be an integral part of future studies to ensure accurate interpretation of results, correct comparison of injury rates, and effective planning of future injury prevention strategies.

Practical Implications

- Exercise and sport science students with appropriate training can reliably collect and report player game exposure and injury number data in large scale sports injury randomised controlled trials.
- Reporting the reliability of data collection processes in relation to exposure and injury incidence data is critical for interpretation and utilisation of collected data.

References

1. Braham R, Finch C, McCrory P. Sports trainers have accurate but incomplete recall of injury details. *Br J of Sports Med* 2003; 37: 561.

2. Yard E, Collins C, Comstock R. A comparison of high school sports injury surveillance data reporting by certified athletic trainers and coaches. *J Athl Training* 2009; 44(6): 645-52.

3. Shrier I, Feldman D, Akakpo H, et al. Discordance in injury reporting between youth-athletes, their parents and coaches. *J Sci Med Sport* 2009; 12: 633-36.

4. Emery C, Meeuwisse W, Hartmann S. Evaluation of risk factors for injury in adolescent soccer: implementation and validation of an injury surveillance system. *Am J Sports Med* 2005;33(12):1882 - 91.

5. Gabbe B. Risk factors for lower limb injuries at the community-level of Australian football [PhD Thesis]. Melbourne: University of Melbourne; 2002.

6. Gabbe B, Finch C, Wasjswelner H, et al. Predictors of lower extremity injuries at the community level of Australian football. *Clin J Sport Med* 2004; 14:56-63.

7. Gabbe B, Finch C, Bennell K, et al. Risk factors for hamstring injuries in community-level Australian Football. *Br J of Sports Med* 2005;39:106-10.

8. Grimmer K, Williams J. Injury in junior Australian rules footballers. *J Sci Med Sport* 2003;6(3):328-38.

9. Braham R, Finch C, McIntosh A, et al. Community level Australian football: a profile of injuries. *J Sci Med Sport* 2004;7(1):96-105.

10. Braham R, Finch C, McCrory P. The incidence of head/neck/orofacial injuries in non-elite Australian football. *J Sci Med Sport* 2004;7(4):451-3.

11. Braham RA, Finch CF. The reliability of team-based primary data collectors for the collection of exposure and protective equipment use data in community sport. *Br J of Sports Med* 2004;38(4):e15.

12. Romiti M, Finch C, Gabbe B. A prospective cohort study of the incidence of injuries among junior Australian football players- evidence for a playing age level effect. *Br J of Sports Med* 2008;42:441-6.

13. Fuller C, Clarke l, Molloy M. Risk of injury associated with rugby union played on artificial turf. *J Sports Sci* 2010; 28: 563-70.

14. Dallalana RJ, Brooks JHM, Kemp SPT, Williams AM. The epidemiology of knee injuries in English professional rugby union. *Am J Sports Med* 2007; 35: 818 - 30.

15. Herbenick M, King J, Altobelli G, et al. Injury patterns in professional arena football. *Am J Sports Med* 2007; 36: 91-8.

16. McMahon K, Nolan T, Bennett C, et al. Australian Rules football injuries in children and adolescents. *Med J Australia* 1993;159:301 - 6.

17. Finch C, Lloyd D, Elliott B. The preventing Australian football injuries with eXercise (PAFIX) study - a group randomised controlled trial. *Inj Prev* 2009;15(e1).

18. Viera A, Garrett J. Understanding interobserver agreement: The kappa statistic. *Fam Med* 2005;37(5):360 - 3.

19. Braham R, Finch C. The reliability of team-based primary data collectors for the collection of exposure and protective equipment use data in community sport. *Br J of Sports Med* 2004;38(4):E15.

20. Lower T. Injury data collection in the rugby codes. *Aust J Sci Med Sport* 1995;27(2):43 - 7.

Table 1: Levels of agreement between the primary data collectors and independent observers.

	% Agreement	ICC (95%CI)	Rating of Agreement
Participation Data			
Number of players	79	0.9 (0.85 – 0.95)	Almost perfect
Injury Data			
No. of injuries	52	0.7(0.48 – 0.85)	Substantial
No. of injuries (adjusted for end of quarter/game)	76	0.8 (0.69 – 0.91)	Substantial
No. of injuries that left the field	69	0.7 (0.52 – 0.86)	Substantial
No. of injuries that left the field (adjusted for injuries collected by both)	100	1.00	Perfect

Table 2: Level of agreement between the number of injuries sustained per game reported by the primary data collectors (PDCs) and the independent observers (IOs).

		PDC					**Total**
		0 injury	1 injury	2 injuries	3 injuries	4 injuries	
IO	0 injury	10	3	0	0	0	**13**
	1 injury	5	7	1	0	0	**13**
	2 injuries	2	3	3	2	0	**10**
	3 injuries	0	1	2	1	1	**5**
	4 injuries	1	0	0	0	0	**1**
Total		**18**	**14**	**6**	**3**	**1**	**42**

Note: Dark shaded diagonal represents the number of times total agreement occurred. The vertical shading represents where the PDCs recorded more injuries and the horizontal shading where the IOs had more injuries.

OFFICE USE ONLY: Database Game ID: _____

PAFIR
PREVENTING AUSTRALIAN FOOTBALL INJURIES THROUGH EXERCISE

OBSERVATIONAL CHECK FORM

Club:_____ **Grade:**_____
Game Day (eg Sat)**:**_____ **Game Date:**_____
Game Venue:_____ **Opposition**:_____
Start Time:_____ **Finish Time:**_____
Final Score (Your team followed by opposition): _____ to _____

Player Jumper Number	**Match Game**									
	Participation (Please tick)				**Match Injury** (Y/N)	**Leave field** (Y/N)	**Match Injury** (Y/N)	**Leave field** (Y/N)	**Match Injury** (Y/N)	**Leave field** (Y/N)
	1st Q	2nd Q	3rd Q	4th Q						

Additional Comments:

Name_____ Signature_____ Date_____